

3-2012

# Performance of Five Models to Predict the Naturalization of Non-Native Woody Plants in Iowa

Emily J. Kapler  
*Iowa State University*

Mark P. Widrlechner  
*United States Department of Agriculture, isumw@iastate.edu*

Philip M. Dixon  
*Iowa State University, pdixon@iastate.edu*

Janette R. Thompson  
*Iowa State University, jrtrt@iastate.edu*

Follow this and additional works at: [http://lib.dr.iastate.edu/nrem\\_pubs](http://lib.dr.iastate.edu/nrem_pubs)



Part of the [Ecology and Evolutionary Biology Commons](#), [Horticulture Commons](#), and the [Natural Resources Management and Policy Commons](#)

The complete bibliographic information for this item can be found at [http://lib.dr.iastate.edu/nrem\\_pubs/2](http://lib.dr.iastate.edu/nrem_pubs/2). For information on how to cite this item, please visit <http://lib.dr.iastate.edu/howtocite.html>.

---

# Performance of Five Models to Predict the Naturalization of Non-Native Woody Plants in Iowa

## Abstract

Use of risk-assessment models that can predict the naturalization and invasion of non-native woody plants is a potentially beneficial approach for protecting human and natural environments. This study validates the power and accuracy of four risk-assessment models previously tested in Iowa, and examines the performance of a new random forest modeling approach. The random forest model was fitted with the same data used to develop the four earlier risk-assessment models. The validation of all five models was based on a new set of 11 naturalizing and 18 non-naturalizing species in Iowa. The fitted random forest model had a high classification rate (92.0%), no biologically significant errors (accepting a plant that has a high risk of naturalizing), and few horticulturally limiting errors (rejecting a plant that has a low risk of naturalizing) (8.7%). Classification rates for validation of all five models ranged from 62.1 to 93.1%. Horticulturally limiting errors for the four models previously developed for Iowa ranged from 11.1 to 38.5%, and biologically significant errors from 4.2 to 18.5%. Because of the small sample size, few classification and error rate results were significantly different from the original tests of the models. Overall, the random forest model shows promise for powerful and accurate risk-assessment, but mixed results for the other models suggest a need for further refinement.

## Keywords

validation, random forest, risk assessment, invasive plants, exotic plants, life history, native range, machine learning, NCRPIS, Statistics

## Disciplines

Ecology and Evolutionary Biology | Horticulture | Natural Resources Management and Policy

## Comments

This article is from *Journal of Environmental Horticulture* 30, no. 1 (March 2012): 35–41.

## Rights

Works produced by employees of the U.S. Government as part of their official duties are not copyrighted within the U.S. The content of this document is not copyrighted.

# Performance of Five Models to Predict the Naturalization of Non-Native Woody Plants in Iowa<sup>1</sup>

Emily J. Kapler<sup>2</sup>, Mark P. Widrlechner<sup>3</sup>, Philip M. Dixon<sup>4</sup> and Janette R. Thompson<sup>5</sup>

*Department of Natural Resource Ecology and Management  
Iowa State University, Ames IA 50011*

## Abstract

Use of risk-assessment models that can predict the naturalization and invasion of non-native woody plants is a potentially beneficial approach for protecting human and natural environments. This study validates the power and accuracy of four risk-assessment models previously tested in Iowa, and examines the performance of a new random forest modeling approach. The random forest model was fitted with the same data used to develop the four earlier risk-assessment models. The validation of all five models was based on a new set of 11 naturalizing and 18 non-naturalizing species in Iowa. The fitted random forest model had a high classification rate (92.0%), no biologically significant errors (accepting a plant that has a high risk of naturalizing), and few horticulturally limiting errors (rejecting a plant that has a low risk of naturalizing) (8.7%). Classification rates for validation of all five models ranged from 62.1 to 93.1%. Horticulturally limiting errors for the four models previously developed for Iowa ranged from 11.1 to 38.5%, and biologically significant errors from 4.2 to 18.5%. Because of the small sample size, few classification and error rate results were significantly different from the original tests of the models. Overall, the random forest model shows promise for powerful and accurate risk-assessment, but mixed results for the other models suggest a need for further refinement.

**Index words:** validation, random forest, risk assessment, invasive plants, exotic plants, life history, native range, machine learning.

## Significance to Nursery Industry

Nursery and landscape professionals introduce many new non-native plants, but sometimes these introductions escape from cultivation, naturalize, and invade. This is a concern to many stakeholders, from members of the nursery industry itself to land managers who must deal with invasive species encroaching on natural areas. As new plants continue to be introduced, there is the possibility of inadvertently ushering in new invasive plants. Given the many benefits of introducing new plants, researchers have worked to develop methods to discern potential invaders from benign introductions through risk-assessment modeling. Plants screened by these models are then recommended for acceptance, rejection, or further study based on plant attributes, such as life-history traits or geographic origin. Errors produced by risk-assessment models represent potential costs, both biologically and horticulturally. This paper focuses on the validation of four existing risk-assessment models for woody plants in Iowa, and the application and validation of a new (and potentially more accurate) 'random forest' modeling technique to predict naturalizing and non-naturalizing plants. Validation, which

represents a 'real world' test of the models, indicates that there is room for improvement in their power and accuracy. The new random forest modeling technique shows promise for the future development of a regional-scale model for the Upper Midwest.

## Introduction

The migration of species across the globe is a natural process, but humans are able to disperse and spread organisms beyond their native ranges much more quickly and extensively than any other species (37). Sometimes this occurs by accident, but with plants, most introductions are deliberately carried out by humans (22). The horticultural industry is widely recognized as a major influence on introduced plants (4, 5, 35). Clearly, most introduced plants are benign and benefit human interests. Sometimes introduced plants will thrive in their new environment and are able to sustain populations without human assistance. A few of these do so well that they begin to aggressively displace native vegetation and alter local ecosystems. This progression from naturalized to invasive is influenced by many factors and may be understood as a continuum (36). While only a small percentage of woody plants will invade outside their native ranges (38), one unintended consequence of introducing new plants is to cause undesirable changes to our landscapes (10, 23, 28, 43) that are costly to manage (27).

In order to prevent undesirable consequences, screening new plants for invasiveness before introduction may be an effective strategy with net bioeconomic benefits (17). This has led to the development of statistical models to evaluate the probability that a non-native plant will naturalize or invade in a new location. Information, such as life-history characteristics of plants that are associated with invasiveness, is typically included; pertinent geographic or climatic variables are often factored in as well (33, 39, 44, 48). Several models have been developed and are based on different kinds of statistical procedures, such as discriminant analyses (32), classification and regression trees (34), and analytic hierarchy processes (25). Some take the form of a scoring system, such

<sup>1</sup>Received for publication September 16, 2011; in revised form December 9, 2011. Journal paper of the Iowa Agriculture and Home Economics Experiment Station, Ames, IA, and supported by Hatch Act, McIntire-Stennis, and State of Iowa funds. We acknowledge additional financial support from USDA-ARS through the Floral and Nursery Crops Research Initiative. We also thank Drs. Michael Dosmann, Jeffery Iles, and Marcel Rejmanek, and two anonymous peer reviewers for their useful critiques of our manuscript, and Matt O'Hearn for technical support. Mention of commercial brand names does not constitute an endorsement of any product by the U.S. Department of Agriculture or cooperating agencies.

<sup>2</sup>Graduate Student, Department of Natural Resource Ecology and Management, Iowa State University.

<sup>3</sup>USDA-ARS Horticulturist and Assistant Professor (Collaborator), Departments of Agronomy and Horticulture, Iowa State University (retired).

<sup>4</sup>Professor, Department of Statistics, Iowa State University.

<sup>5</sup>Professor, Department of Natural Resource Ecology and Management, Iowa State University.

as the Australian Weed Risk Assessment (WRA) (26), and others are decision trees (34, 48).

Existing models usually assign a plant to one of three screening outcomes: 'accept' if the plant is at low risk of becoming invasive, 'reject' if the plant is at high risk of becoming invasive, and 'further analysis' where the model is unable to make a clear determination. Power and accuracy of the models can be assessed by testing known invaders and non-invaders (14, 20, 49). Classification rates (which determine the 'power' associated with the models) are based on the proportion of species a model classifies, and should ideally be high, given the time and expense of reassessing 'further analysis' outcomes (44). Models may also produce two types of errors (which reflect their 'accuracy'): 1) false positives, or horticulturally limiting errors which incorrectly reject a plant that actually has a low risk of becoming invasive, and 2) false negatives, or biologically significant errors which incorrectly accept a plant that has a high risk of becoming invasive (48). Given the potential costs associated with these errors, researchers continue to test, validate, and improve risk-assessment models to minimize these problems.

One way of improving models is to tailor them to more specific geographic regions. Risk-assessment models for woody plants, in particular, may benefit from this approach, because of the importance of local climatic and edaphic conditions in influencing woody-plant survival (45, 46). Widrlechner and Iles (47) established a list of 100 non-native woody plants cultivated in Iowa that were either naturalized (28 species) or non-naturalized (72 species). This plant list was used to test an existing continental-scale model (34) and generate three new models to predict the likelihood that these species would escape from cultivation and potentially become invasive in Iowa (48). Model validation can be done internally during model development, but can also be done externally by testing a new data set from the same region or from a similar region. The models from Widrlechner et al. (48) were externally validated by using independent datasets for non-native woody plants from the Chicago region with mixed results (49).

A second way to improve the overall performance of risk-assessment models is to apply different statistical techniques that may yield better power and accuracy. Classification and regression tree (CART) approaches have previously been used to develop risk-assessment models (e.g., 34) with some success, but they have some inherent limitations. CART trees differentiate species within a data set by using a series of dichotomous branches based on classification rules derived from a training data set (e.g., an initial list of naturalizing and non-naturalizing species) (13). Each subsequent decision node (which is based on a classification rule) is developed with a progressively smaller sample size. This makes the classification rules for nodes further down the tree very sensitive to small changes in the training data set, generating high variance (24). Another statistical approach, random forest modeling, can reduce this variance by averaging many classification trees based on small perturbations of the original data (1). In this way, the small sample sizes used to determine terminal classification rules become less of an issue, because the list of species used to make this rule is subject to additional randomization. Random forest modeling also includes a step to reduce positive correlations among predictions to further reduce variance and allows for assessment of variable importance within the model.

Random forest models have been documented as more robust and more accurate than CART models (12, 13). For example, Cutler et al. (3) tested four different classification methods (including CART) to predict the presence of four invasive plant species and found that a random forest approach outperformed the other methods in most accuracy measures. Additional studies by Williams et al. (50), Kampichler et al. (15), and Keller et al. (18) also suggest that a random forest approach may be valuable for developing risk-assessment models to predict the naturalization of non-native woody plants.

In this paper, we report on two avenues of research. First, motivated by mixed results from external validation of the Iowa models (48) when tested with Chicago-region datasets (49), we conducted an evaluation of the Iowa models with a new dataset that matched the region of their development. Second, we tested the performance of the random forest approach for use as a risk-assessment model to predict naturalization of woody plants in Iowa.

## Materials and Methods

We began by generating a list of non-native woody plant species cultivated in Iowa, not included in Widrlechner et al. (48), that could be clearly assigned to categories either as naturalizing or non-naturalizing in the study area. New naturalizing species were determined by examining herbarium vouchers that had not been collected or available when the previous list was made (47). At least two distinct voucher records from Iowa indicating reproduction outside of cultivation needed to be present for a species to be considered as naturalized. Additional non-naturalizing species were suggested by the authors and Jeffery Iles; herbarium records were checked for these species to confirm that they had no record of naturalization. If a species had only one record suggesting naturalization, it was left out of the study. Both lists were then examined for accuracy and completeness by individuals experienced with the Iowa flora (Deborah Lewis, Jimmie Thompson, Cathy Mabry McMullen, and Mark Vitosh). This process resulted in a list of 29 additional non-native woody species cultivated in Iowa. Of these, 11 species have naturalized and 18 have no evidence of naturalization in Iowa.

For each of these 29 species, data on life-history characteristics (Table 1) and native ranges required by the models were compiled. These data were obtained from previous work and several published and online sources (6, 31, 41, 49) with additional review by the authors and professionals with experience cultivating these plants. The native ranges of the 29 species across 278 geographic subdivisions were used to calculate geographic-risk values (as per 48). Native range data were primarily obtained from the USDA-ARS Germplasm Resources Information Network database (42) and previous data from the Chicago study (49), with supplementation from published floras (8, 19, 40). Geographic risk values (G-values) for these species were calculated on the basis of the proportion of species native to a geographic subdivision that have naturalized in Iowa, as described by Widrlechner et al. (48). These proportions were already determined for nearly all geographic subdivisions in our current study. In those few cases (approximately 7% of 1000 data cells) where we found a plant occurring in a geographic subdivision that had not been treated by Widrlechner et al. (48), values based on neighboring or similar subdivisions were used if available, or the subdivision was considered as a missing data cell.

Table 1. Characteristics of 29 new non-native woody landscape plants in Iowa used to test models to assess the risk of naturalization in Iowa (see (32, 48) for more information on these characteristics).

Species	Naturalized	G*	Invades outside North America	Requires germination pretreatment	Group invasive in North America	Quick maturity	Sterile hybrid	Quick vegetative spread	North American native	Evergreen foliage	Fleshy, bird-dispersed fruits
<i>Acer palmatum</i> Thunb.	N	0.315	Y	Y	N	N	N	N	N	N	N
<i>Amelanchier canadensis</i> (L.) Medik.	N	0.055	N	Y	Y	Y	N	N	Y	N	Y
<i>Buddleia davidii</i> Franch.	N	0.260	Y	N	N	Y	N	N	N	N	N
<i>Buxus sempervirens</i> L.	N	0.443	N	Y	N	N	N	N	N	Y	N
<i>Castanea dentata</i> (Marshall) Borkh.	N	0.071	N	Y	N	N	N	N	Y	N	N
<i>Clematis terniflora</i> DC.	N	0.373	N	Y	N	Y	N	Y	N	N	N
<i>Cotoneaster divaricatus</i> Rehder & E.H. Wilson	N	0.339	Y	Y	Y	Y	N	N	N	N	Y
<i>Cotoneaster horizontalis</i> Decne.	N	0.257	Y	Y	Y	N	N	N	N	N	Y
<i>Fothergilla gardenii</i> L.	N	0.110	N	Y	N	N	N	N	Y	N	Y
<i>Hedera helix</i> L.	N	0.504	Y	Y	N	N	N	Y	N	Y	Y
<i>Picea abies</i> (L.) H. Karst	N	0.491	N	N	N	N	N	N	N	Y	N
<i>Pinus sylvestris</i> L.	N	0.444	Y	N	N	N	N	N	N	Y	N
<i>Platycladus orientalis</i> (L.) Franco	N	0.367	N	Y	N	N	N	N	N	Y	N
<i>Prunus cerasifera</i> Ehrh.	N	0.430	Y	Y	Y	Y	N	N	N	N	Y
<i>Salix caprea</i> L.	N	0.452	Y	N	Y	Y	N	N	N	N	N
<i>Syringa pubescens</i> Turez. subsp. <i>patula</i> (Pall.) M.C. Chang & X.L. Chen	N	0.421	N	N	Y	N	N	N	N	N	N
<i>Taxodium distichum</i> (L.) Rich.	N	0.152	N	Y	N	N	N	N	Y	N	N
<i>Wisteria floribunda</i> (Willd.) DC.	N	0.339	Y	N	Y	N	N	Y	N	N	N
<i>Acer platanoides</i> L.	Y	0.485	Y	Y	N	N	N	N	N	N	N
<i>Celastrus orbiculatus</i> Thunb.	Y	0.367	Y	Y	N	Y	N	Y	N	N	Y
<i>Fraxinus alnus</i> Mill.	Y	0.439	N	Y	N	N	N	N	N	N	Y
<i>Ligustrum obtusifolium</i> Siebold & Zucc.	Y	0.359	Y	Y	Y	Y	N	N	N	N	Y
<i>Ligustrum vulgare</i> L.	Y	0.453	Y	Y	Y	Y	N	N	N	N	Y
<i>Lonicera japonica</i> Thunb.	Y	0.318	Y	Y	Y	Y	N	Y	N	SEMI	Y
<i>Lonicera morrowii</i> A. Gray	Y	0.343	N	Y	Y	Y	N	N	N	N	Y
<i>Lonicera sempervirens</i> L.	Y	0.142	N	Y	Y	Y	N	N	Y	N	Y
<i>Rhamnus davurica</i> Pall.	Y	0.469	N	Y	Y	N	N	N	N	N	Y
<i>Rhamnus utilis</i> Decne.	Y	0.288	N	Y	Y	N	N	N	N	N	Y
<i>Rubus caesius</i> L.	Y	0.484	Y	Y	Y	Y	N	Y	N	N	Y

\*Geographic-risk value, a ratio that can vary between zero and one, is the average proportion of naturalizing to non-naturalizing species (relative to Iowa) based on the many geographic subdivisions that encompass the species' native range.

These data were collected and reviewed and then the four risk assessment models described in detail by Widrechner et al. (48, 49) were applied to the 29 new species. These models included Reichard & Hamilton's 'continental decision tree' (34) and three additional models developed specifically for Iowa (48): 1) the 'modified decision tree' which adds ten steps to the continental decision tree, 2) the 'decision tree/matrix model' which focuses on reevaluating the 'further analysis' species produced by the continental decision tree, and 3) the 'CART model' developed specifically for the original Iowa data set and based on a classification and regression tree (CART).

In addition, a new random forest model was created based on the dataset of 100 species from the original Iowa study (47, 48). A random forest (1, 3) is an extension of a CART model. As noted in the introduction, a CART model partitions data into smaller and smaller subsets, so its predictions can be quite variable.

In detail, the random forest algorithm includes:

1. Drawing a non-parametric bootstrap sample (7) of the observations (in our case, an observation is a woody plant species). Some observations are omitted from the bootstrap sample, some observations occur once, and others are repeated multiple times.
2. Constructing a CART model based on the bootstrap data. At each potential split, a randomly selected subset of the variables is evaluated to define a split. This random selection of variables reduces the positive correlation among predictions and improves the precision of the prediction.
3. Calculating the probability of naturalization for each observation in the bootstrap sample.
4. Repeating steps 1 through 3 for 1000 bootstrap samples.
5. Calculating the average probability of naturalization for an observation by averaging predictions for that observation in all CART trees.

A fitted random forest model was created based on the original 100 Iowa species generated from 1000 CART trees. We fit random forests using the randomForest and helper functions in the randomForest package (21) in the R program, version 2.12.2 (30). The probability of not naturalizing was set equal to 0.72, the proportion of species without evidence of naturalizing in the original 100-species data set for Iowa.

The fitted random forest was used to predict the probability of naturalization for each of the 100 species in the training data set (the species list used to develop the model) and for each of the 29 new species. The classification of species as 'accept', 'reject', or 'further analysis' was based on the predicted probability of naturalization. Comparing the predicted probabilities to the observed status of each of the 100 species in the training data set supported the following classification rule:

- If the predicted probability is  $< 0.12$ , then classify as 'accept';
- If the predicted probability is  $\geq 0.28$ , then classify as 'reject'; and
- If the predicted probability is between 0.12 and 0.28, classify as 'further analysis'.

The power and accuracy of each model were assessed in the following manner. First, we examined the 'classification rate,' or the proportion of species successfully assigned 'ac-

cept' or 'reject' by the models. We also assessed two types of errors, the 'horticulturally limiting error' and 'biologically significant error', expressed as the proportion of error to the total number of classified species (as per 48, 49).

The statistical significance of differences in classification rates among models was assessed by reducing the classification of species to two groups: successfully classified or further analysis. The null hypothesis that all five models had the same probability of successfully classifying a species was tested with a Cochran-Mantel-Haenszel test for stratified categorical data (9), with each species considered a unique stratum. This statistical test accounts for species-species differences in ease of classification. When the Cochran-Mantel-Haenszel test was significant, individual models were compared to the average performance to identify which models performed better or worse than average. Because each stratum had at most five observations (one per method), p-values for all statistical tests were computed by randomization within strata, using 999 permuted data sets. Variable importance for the random forest model was assessed by measuring the total decrease in Gini impurity for splits involving each variable (12).

The statistical significance of differences in horticulturally limiting errors and biologically significant errors was assessed by reducing the classification to 'accept' or 'reject' and treating all 'further analysis' results as missing values. The random alteration of each species only permuted 'accept' or 'reject' values to the classified observations, i.e. the missing values were not permuted. This approach compares the probability of a biological or horticulturally limiting error among models when the method classified a species. Statistical significance of the differences in classification and error rates between old and new data sets was assessed with the Fisher exact test for  $2 \times 2$  tables (9). All statistical tests of differences and variable importance were done with R statistical software (30).

## Results and Discussion

*Performance of the four original models on 29 new species.* The set of four models tested previously (48) had variable performance when applied to the new set of 29 Iowa species. Classification rates ranged from 62.1 to 93.1% (Table 2), which is comparable to classification rates for other types of models (11, 14, 20). Comparing classification rates for the 29 new species to the original 100 species, the continental decision tree performed better for the new species ( $P < 0.01$ ) and the CART model performed worse ( $P < 0.05$ ); other classification rates did not differ significantly.

Two of the models are based on modifications to the continental decision tree. The refinements of the modified decision tree were designed to focus on the branch of that decision tree that produced the most errors and 'further analysis' outcomes (48). Because nine out of the ten species producing horticulturally limiting errors in the new set of 29 Iowa species came from the branch targeted by the modified decision tree model, this new test set underscores the importance of this step. However, its ability to produce improvements was mixed. While there was a reduction in horticulturally limiting errors, two species generated biologically significant errors (Table 2). The second model based on the continental decision tree — the decision tree/matrix model — focused on reanalyzing 'further analysis' species. Given the high initial classification rate of the continental decision tree for the 29 new Iowa species, there was little



Table 2. Summary of classification and error rates for five risk-assessment models by data set.

Model	Classification rate (%)	Biologically significant error rate (%)	Horticulturally limiting error rate (%)
Continental decision tree			
Original 100 Iowa species	65.0	3.1	16.9
New 29 Iowa species	89.7	11.5	38.5
Modified decision tree			
Original 100 Iowa species	90.0	3.3	13.3
New 29 Iowa species	93.1	18.5	29.6
Decision tree/matrix model			
Original 100 Iowa species	85.0	3.5	16.4
New 29 Iowa species	93.1	14.8	37.0
CART model			
Original 100 Iowa species	81.0	2.5	3.7
New 29 Iowa species	62.1	16.7	11.1
Random forest model			
Original 100 Iowa species	92.0	0.0	8.7
New 29 Iowa species	82.8	4.2	29.2

room for improvement. One species (*Lonicera sempervirens*) was treated differently between these models, and it became a biologically significant error. The CART model, which is not related to the continental decision tree, had a much lower classification rate but, to its credit, it displayed the best (lowest) horticulturally limiting error rate (Table 2). It also had a higher biologically significant error rate, though it misclassified the same number of species (three) as did the continental decision tree. The biologically significant error rate for CART is relatively higher in this case because there are more 'further analysis' outcomes, decreasing the denominator used to calculate error rates.

Differences in error rates between the original 100 species and the 29 new species were not statistically significant (probably due to the small sample size of the new species data) with one exception: the horticulturally limiting error rate for the continental decision tree was greater for the new species tested ( $P < 0.02$ ). Both types of error rates for the four original models were, however, higher for the new 29 species than for the original 100 species, ranging from 11.5 to 18.5% for biologically significant errors and from 11.1 to 38.5% for horticulturally limiting errors (Table 2). They are also higher than error rates reported in the Chicago study (49) or for many tests of other risk-assessment models, such as the Australian WRA (see 11 for a meta-analysis). Similar to results reported by Widrlechner et al. (49), the CART model had the lowest horticulturally limiting error rate of the four, which is encouraging given that many other risk-assessment models generate few biologically significant errors at the expense of more horticulturally limiting errors.

The high error rates overall are not surprising given the nature of the data set. These 29 species represent, in many respects, a 'real world' test in that they do not conform to the 0.28 ratio of naturalizing species to non-naturalizing species under which three of the four models were developed (48). Models should ideally be robust enough to perform well under deviations from this ratio, such as the 0.38 ratio that we observed for the 29 new Iowa species. There are also idiosyncrasies that arise from the list of plants themselves. This pool of naturalizing species is different in some important ways. Since these species are based on newer records of naturalization, there are fewer 'major invaders' of Iowa than were included in the list used to develop the models. Krivánek and Pyšek (20) have suggested that woody plant risk-assessment models are generally better at pinpointing

strongly invasive species than at sorting out those which have only begun to naturalize.

Certain species tended to produce errors across all four of the models. In each of the models, *Frangula alnus* and *Rhamnus utilis* generated biologically significant errors; *Rhamnus davurica*, *Acer platanoides* and *Lonicera sempervirens* were other common sources of errors. Two species also generated horticulturally limiting errors in all four models: *Prunus cerasifera* and *Salix caprea*. Other common horticulturally limiting errors, generated by three models, were *Buddleja davidii*, *Clematis terniflora*, *Cotoneaster divaricatus*, *Cotoneaster horizontalis*, and *Hedera helix* [each of these four species also produced errors in the Chicago region (49)]. There is always the possibility that species presently categorized as horticulturally limiting errors will naturalize in the future, due to the considerable lag-time between introduction and naturalization for woody plants (2, 29). Two of these species (*Buddleja davidii*, *Clematis terniflora*) are known to have naturalized in northern Missouri and could conceivably do the same in Iowa in the coming decades. Overall, the performance of these four models on the test set of 29 new Iowa species resulted in disappointing error rates, highlighting the need for continued model development.

*Performance of the random forest model.* The fitted random forest model, which is the product of 1000 decision trees trained on the original 100 Iowa species, performed well overall. Classification rates (Table 2) were significantly better than the average rate for all other models ( $P = 0.002$ ). The biologically significant error rate was zero and also significantly better than the average of all other models ( $P = 0.018$ ). At the same time, the fitted random forest model was able to discern non-naturalizing species better than three of the other models ( $P = 0.092$ ), but of the five models, the CART model produced the fewest horticulturally limiting errors ( $P = 0.016$ ). Although the fitted random forest model was not the best for horticulturally limiting errors, it still performed well overall, confirming the strength of random forest modeling when applied to risk-assessment for non-native woody plants. It also performs well compared to the classification and error rates of other risk-assessment models in the literature (i.e., 11, 14, 20).

Validation of the fitted random forest model based on the 29 new Iowa species was somewhat less impressive, but still promising. The classification rate dropped, but was

not different from the average of the other models tested on the same set of species ( $P = 1.00$ ). Of the five models, the fitted random forest model had a relatively low biologically significant error rate (Table 2), but it was not significantly different ( $P = 0.646$ ) from the others (again, perhaps because of the small sample size). All five models had low biologically significant error rates (Table 2) for the 29 new species. Again, the random forest model was not statistically significantly different from the other models ( $P = 0.65$ ). The random forest model had the second best horticulturally limiting error rate for the 29 new species, although the evidence of a difference is weak ( $P = 0.092$ ).

We assessed the relative importance of each variable in the predictions made by the fitted random forest model. Variable importance was determined by the average amount that each variable reduced the uncertainty in the predicted probability of naturalization. Geographic-risk values and quick maturation were the two most important characteristics for determining the ability of a plant to naturalize in Iowa, followed by whether it is invasive outside North America and whether it has fleshy, bird-dispersed fruits (Fig. 1). The importance of these variables in the random forest model may also help explain some of the strengths of the CART model in this and in previous studies (48, 49), since the CART model includes only G-values, quick maturity, and fleshy, bird-dispersed fruits as predictive variables. The variable importance results also strongly suggest that correctness in determining these four traits for any non-native plant to be introduced to Iowa is of greatest importance for assessment accuracy.

**General conclusions.** The relatively high classification rate (82.8%) of the random forest model indicates that it may be a promising approach for predicting naturalization of non-native woody plants. It does, however, have some drawbacks that may limit its use by those who wish to screen non-native plants for invasiveness (e.g., personnel associated with public gardens, arboreta, or nurseries). Because a fitted random forest model is the product of many decision trees, it cannot be presented as a single, easy-to-understand diagram like the other four models. It becomes a 'black box' where data go in and recommendations mysteriously emerge, and it requires specific technical skills to use, including familiarity

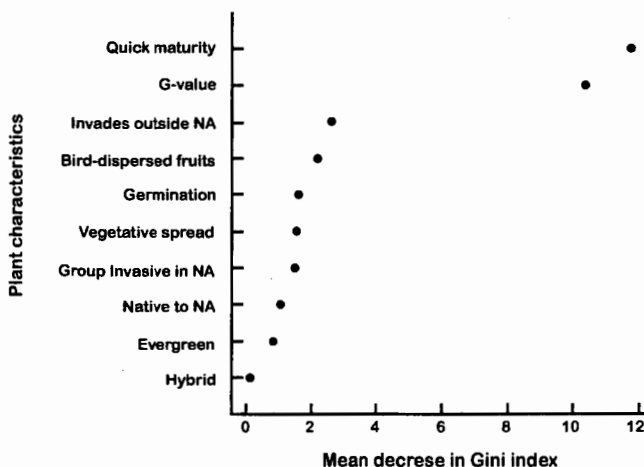


Fig. 1. Variable importance in the random forest model based on 100 Iowa species (48).

with statistical software such as R (13, 24). As such, application of the random forest approach might require external technical support and funding to conduct the analyses, but we are developing graphical approximations to simplify use of the random forest. It is here where the other models, in spite of their mixed performance during validation, have the advantage of easier use for testing individual species.

We know from surveys focused on risk-assessment models in Iowa that stakeholders (including conservation professionals, master gardeners, professional horticulturists, and woodland landowners) prefer low biologically significant error rates, and that based on stakeholders' median values they believe such errors should not exceed 10% (16). This suggests that the random forest model would be an acceptable choice for stakeholders, based on the external validation of the 29 new species. However, validation of the random forest model exceeded stakeholder preferences for a 20% upper limit for horticulturally limiting errors (16); only the CART model fit this limit for the 29 new species (Table 2). Horticulturally limiting error rates always need to be interpreted with care, as some apparent errors may forecast future naturalization events. Even if some of these errors may be explained by idiosyncrasies in the species list or the likelihood of future naturalization, there is still a need to reduce this type of error in the random forest model. To this end, we intend to complete additional validations of these risk-assessment models on two additional data sets from the Upper Midwestern United States, one from northern Missouri and the other from southern Minnesota. Our ultimate goal is to produce a regional model to predict the naturalization of non-native woody plants that is more accurate, powerful, and easy to use than models currently available.

## Literature Cited

- Breiman, L. 2001. Random forests. *Mach. Learn.* 45:5–32.
- Crooks, J.A. 2005. Lag times and exotic species: The ecology and management of biological invasions in slow-motion. *EcoScience* 12:316–329.
- Cutler, D.R., T.C. Edwards Jr., K.H. Beard, A. Cutler, K.T. Hess, J. Gibson, and J.J. Lawler. 2007. Random forests for classification in ecology. *Ecology* 88:2783–2792.
- Dawson, W., A.S. Mdolwa, D.F.R.P. Burselm, and P.E. Hulme. 2008. Assessing the risks of plant invasions arising from collections in tropical botanical gardens. *Biodivers. Conserv.* 17:1979–1995.
- Dehnen-Schmutz, K., J. Touza, C. Perrings, and M. Williamson. 2007. The horticultural trade and ornamental plant invasions in Britain. *Conserv. Biol.* 21:224–231.
- Dirr, M.A. 1998. *Manual of Woody Landscape Plants*, 5th ed. Stipes, Champaign, IL.
- Dixon, P.M. 2002. Bootstrap resampling. *Encyclopedia of Environmetrics* 1:212–220.
- eFloras 2010. *Flora of China* online volumes. Accessed 2010. [http://www.efloras.org/flora\\_page.aspx?flora\\_id=2](http://www.efloras.org/flora_page.aspx?flora_id=2).
- Fleiss, J.L. 1981. *Statistical Methods for Rates and Proportions*, 2nd ed. Wiley, New York, NY.
- Gaertner, M., A. Den Breeyen, C. Hui, and D.M. Richardson. 2009. Impacts of alien plant invasions on species richness in Mediterranean-type ecosystems: A meta-analysis. *Prog. Phys. Geog.* 33:319–338.
- Gordon, D.R., D.A. Onderdonk, A.M. Fox, and R.K. Stocker. 2008. Consistent accuracy of the Australian weed risk assessment system across varied geographies. *Divers. Distrib.* 14:234–242.
- Hastie, T., R. Tibshirani, and J. Friedman. 2009. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd Ed. Springer, New York, NY.



13. Jarošík, V. 2011. Cart and related methods. p. 104–108. *In*: D. Simberloff and M. Rejmánek (Eds.). *Encyclopedia of Biological Invasions*. University of California Press, Berkeley, CA.
14. Jefferson, L., K. Havens, and J. Ault. 2004. Implementing invasive screening procedures: The Chicago Botanic Garden model. *Weed Technol.* 18:1434–1440.
15. Kampichler, C., R. Wieland, S. Calmé, H. Weissenberger, and S. Arriaga-Weiss. 2010. Classification in conservation biology: A comparison of five machine-learning methods. *Ecol. Inform.* 5:441–450.
16. Kapler, E.J. 2011. Risk analysis for invasive plants in Iowa: Development of risk-assessment models and the perceptions of stakeholders. MS Thesis. Iowa State University, Ames, IA.
17. Keller, R.P., D. Lodge, and D.C. Finnoff. 2007. Risk assessment for invasive species produces net bioeconomic benefits. *Proc. Natl. Acad. Sci. (USA)* 104:203–207.
18. Keller, R.P., D. Kocov, and S. Džeroski. 2011. Trait-based risk assessment for invasive species: High performance across diverse taxonomic groups, geographic ranges and machine learning/statistical tools. *Divers. Distrib.* 17:451–461.
19. Komarov, V.L. et al. (Eds.) 1934–1964. *Flora of the USSR*, 30 vols. Academy of Sciences of the USSR, Moscow and Leningrad.
20. Křivánek, M. and P. Pyšek. 2006. Predicting invasions by woody species in a temperate zone: A test of three risk assessment schemes in the Czech Republic (Central Europe). *Divers. Distrib.* 12:319–327.
21. Liaw, A. and M. Wiener. 2002. Classification and regression by random forest. *R News* 2:18–22.
22. Mack, R.N. and M. Erneberg. 2002. The United States naturalized flora: Largely the product of deliberate introductions. *Ann. Mo. Bot. Gard.* 89:176–189.
23. Mack, R.N., D. Simberloff, W.M. Lonsdale, H. Evans, M. Clout, and F.A. Bazzaz. 2000. Biotic invasions: Causes, epidemiology, global consequences, and control. *Ecol. Appl.* 10:698–710.
24. Olden, J.D., J.J. Lawler, and N.L. Poff. 2008. Machine learning methods without tears: A primer for ecologists. *Q. Rev. Biol.* 83:171–193.
25. Ou, L., C. Lu, and D.K. O'Toole. 2008. A risk assessment system for alien plant bio-invasion in Xiamen, China. *J. Environ. Sci.* 20:989–997.
26. Pheloung, P.C. 2001. Weed risk assessment for plant introductions to Australia. p. 83–92. *In*: R.H. Groves, F.D. Panetta, and J.G. Virtue (Eds.). *Weed Risk Assessment*. CSIRO Publishing, Collingwood, Australia.
27. Pimentel, D. 2011. *Biological Invasions: Economic and Environmental Costs of Alien Plant, Animal, and Microbe Species*, 2<sup>nd</sup> Ed. CRC Press, New York, NY.
28. Powell, K.I., J.M. Chase, and T.M. Knight. 2011. A synthesis of plant invasion effects on biodiversity across spatial scales. *Am. J. Bot.* 98:539–548.
29. Pyšek, P. and K. Prach. 1993. Plant invasions and the role of riparian habitats: A comparison of four species alien to central Europe. *J. Biogeogr.* 20:413–420.
30. R Development Core Team. 2011. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/>.
31. Randall, R. 2003. Rod Randall's big weed list. Accessed 2010. <http://www.invasive.org/gist/biglist.html>.
32. Reichard, S. 1994. Assessing the potential of invasiveness in woody plants introduced to North America. Dissertation. University of Washington, Seattle, WA.
33. Reichard, S. 2001. The search for patterns that enable prediction of invasion. p. 10–19. *In*: R.H. Groves, F.D. Panetta, and J.G. Virtue (Eds.). *Weed Risk Assessment*. CSIRO Publishing, Collingwood, Australia.
34. Reichard, S.H. and C.W. Hamilton. 1997. Predicting invasions of woody plants introduced into North America. *Conserv. Biol.* 11:193–203.
35. Reichard, S.H. and P. White. 2001. Horticulture as a pathway of invasive plant introductions in the United States. *BioScience* 51:103–113.
36. Rejmánek, M. 2011. Invasiveness. p. 379–385. *In*: D. Simberloff and M. Rejmánek (Eds.). *Encyclopedia of Biological Invasions*. University of California Press, Berkeley, CA.
37. Ricciardi, A. 2007. Are modern biological invasions an unprecedented form of global change? *Conserv. Biol.* 21:329–336.
38. Richardson, D.M. and M. Rejmánek. 2011. Trees and shrubs as invasive alien species — A global review. *Divers. Distrib.* 17:788–809.
39. Richardson, D.M. and W. Thuillier. 2007. Home away from home: Objective mapping of high-risk source areas for plant introductions. *Divers. Distrib.* 13:299–312.
40. Tutin, T.G. et al. (Eds.) 1964–1994. *Flora Europaea*, 5 vols. Cambridge University Press, Cambridge.
41. U.S. Department of Agriculture, Forest Service. 2008. *The Woody Plant Seed Manual*. Accessed 2010. [http://www.nsl.fs.fed.us/nsl\\_wpsm.html](http://www.nsl.fs.fed.us/nsl_wpsm.html).
42. U.S. Department of Agriculture, Agricultural Research Service. 2010. Germplasm Resources Information Network database. <http://www.ars-grin.gov/npgs/>.
43. Vilá, M., J.L. Espinar, M. Hejda, P.E. Hulme, V. Jarošík, J.L. Maron, J. Pergl, U. Schaffner, Y. Sun, and P. Pyšek. 2011. Ecological impacts of invasive alien plant: A meta-analysis of their effects on species, communities and ecosystems. *Ecol. Lett.* 14:702–708.
44. White, P.S. and A.E. Schwartz. 1998. Where do we go from here? The challenges of risk assessment for invasive plants. *Weed Technol.* 12:744–751.
45. Widrlechner, M.P. 1994. Environmental analogs in the search for stress-tolerant landscape plants. *J. Arboric.* 20:114–119.
46. Widrlechner, M.P. 2001. The role of environmental analogs in identifying potentially woody plants in Iowa. *J. Iowa Acad. Sci.* 108:158–165.
47. Widrlechner, M.P. and J.K. Iles. 2002. A geographic assessment of the risk of naturalization of non-native woody plants in Iowa. *J. Environ. Hort.* 20:47–56.
48. Widrlechner, M.P., J.R. Thompson, J.K. Iles, and P.M. Dixon. 2004. Models for predicting the risk of naturalization of non-native woody plants in Iowa. *J. Environ. Hort.* 22:23–31.
49. Widrlechner, M.P., J.R. Thompson, E.J. Kapler, K. Kordecki, P.M. Dixon, and G. Gates. 2009. A test of four models to predict the risk of naturalization of non-native woody plants in the Chicago Region. *J. Environ. Hort.* 27:241–250.
50. Williams, J.N., C. Seo, J. Thorne, J.K. Nelson, S. Erwin, J.M. O'Brien, and M.W. Schwartz. 2009. Using species distribution models to predict new occurrences for rare plants. *Divers. Distrib.* 15:565–576.